# HOURLY SOLAR ENERGY PATTERNS USING DATA ANALYSIS: A REVIEW

Mumtaz Anwar Khan

Research Scholar, Computer Science and Technology, Sam Global University

*Abstract*

*Solar energy has emerged as one of the top renewable energy sources, mainly due to its availability all over the world, lower installation costs, and good integration with modern power systems. Nevertheless, its natural intermittency and hourly variability pose operational difficulties not only for PV generation but also for grid control. The traditional method of resource evaluation, which is based on daily and monthly averages to a large extent, is unable to reflect short-term variations that are caused by cloud movements, temperature changes, and atmospheric fluctuations, thus creating uncertainty in forecasting, inverter performance, and battery usage. The advances in high-resolution sensors, satellite radiation mapping, and AI-based solar energy prediction models have made it possible to identify hourly solar energy patterns in detail and thereby to facilitate better control, accuracy in forecasting, and energy-system optimization. The paper reviews and puts together the trends and issues in the areas of irradiance measurements, environmental influence studies, forecasting techniques, and operational PV performance highlighting the key analytical trends, limitations, and future research directions in data-driven hourly solar analysis.*

*Keywords: Solar irradiance, hourly analysis, performance ratio, photovoltaic systems, inverter efficiency, forecasting, environmental impact, data analytics.*

## I.    INTRODUCTION

Hourly solar analysis became a very popular research topic area when scientists realized that diurnal patterns consist of different phases like morning ramp-up, midday thermal losses, and evening reduction. The energy losses at mid-day due to temperature rise, which in turn lowers semiconductor carrier mobility and conversion efficiency even under high irradiance, are hidden by traditional daily averages. The same applies to cloud dynamics that can cause very fast drops in irradiance even on days that are considered sunny. Since these abrupt changes often transpire within less than one hour, hourly resolution becomes imperative for operational forecasting, inverter scheduling, and storage planning. Jiménez-Pérez and Mora-López have shown that using advanced pattern recognition gives better hourly irradiance estimation as compared to the conventional statistical methods under different sky conditions [1].

One of the key advances in the area of hourly solar analytics is the increasing availability of different sensing technologies. Newer PV plants have installed several sensors like pyranometers, irradiance sensors, thermal probes, IoT devices, and inverter telemetry that work together to create large data sets on environmental variables and power outputs with timestamps. In addition to ground-based sensors, numerical-weather-prediction (NWP) models and satellite irradiance retrieval also provide global measurements with the necessary time resolution, making it possible to do forecasting in a wide range of climates. Vijayakumar et al. have demonstrated the significance of very short-term radiation analysis for the understanding of rapid atmospheric transitions, showing that even an hour's worth of fluctuations has a meaningful influence on power generation [2].

With the increase of PV penetration, the hourly behavior becomes more important not only for the plant operations but also for the grid operators. Forecasting hourly for reserve scheduling, load shifting, and ramp-rate management is a necessity for grid operators. Storage systems rely on the optimization of their charging and discharging based on the predicted solar output which is time-sensitive. Distributed PV installations make the situation even more complicated, as the production at the household level is greatly determined by local micro-climates and shading, leading to highly diverse generation profiles. All these factors call for the by far most detailed hourly assessment of irradiance, temperature and inverter performance.

Solar dynamics by the hour are still very much relevant in the case of renewable energies which are coming up. The tropical and desert regions having already mentioned high irradiance are, at the same time, also dealing with the problems of high temperatures, shading, dust accumulation and changes in monsoon seasons. The environment's characteristics have a great impact on the reliability and life span of PV systems. Therefore, the universal applicability of hourly solar studies is acknowledged; however, localized modeling is still a must for accurate forecasting in particular climate zones.

On the other hand, hourly analysis is also very much informative regarding the long-term operational degradation. The cycling of temperature, shading difficulty, and load variations are all part of what contributes to the stress on material and efficiency loss over the life of an inverter, hence making hourly data indispensable for reliability assessment. The combination of advanced forecasting and operational monitoring opens up the possibility of implementing predictive maintenance based on data-driven insights.

In addition, the incorporation of photovoltaic technology with smart grids, the Internet of Things, and advanced analytics is uplifting the importance of hourly modeling in energy-system control. The application of machine learning is heavy in areas such as short-term forecasting, anomaly detection, and operational optimization. According to Ma et al., intelligent energy networks are dependent on real-time analytics to ensure the reliability of the grid and to control distributed assets under renewable energy penetration [3].

In a nutshell, the process of hourly solar analysis has changed from simple irradiance measurement to a complete data-driven method that involves environmental sensing, statistical modeling, deep learning, inverter performance characterization, and smart-grid applications. The worldwide increase in solar penetration will not only demand the management of hourly variability but also determine the operational reliability and economic viability of solar energy systems.

## II. BACKGROUND AND THEORY

The solar radiation that falls on the Earth's surface comes from the Sun and goes through different processes in the atmosphere before it reaches the ground. The processes are scattering, absorption, and interactions with clouds, and the final result is that only a part of the solar energy reaching the Earth is usable and that depends mainly on the irradiance hitting the photovoltaic panel surface. Irradiance is affected by the solar energy location on Earth, the season, the time of day, and the turbidity of the atmosphere. Even though astronomical cycles are predictable, the atmosphere causes great variability. This is the reason why solar resource forecasting is particularly difficult when dealing with hourly data.

Radiation at Earth's surface consists of primarily two parts: direct normal irradiance and diffuse irradiance. Under clear sky conditions, direct radiation is the main part, while during cloudy or aerosol-heavy conditions; the diffuse part is the strongest. Thus, hourly radiation patterns are very different for different sky conditions even if the total for the day is the same. These differences have a great impact because photovoltaic (PV) modules are capable of converting direct and diffuse radiation differently according to their orientation and tracking geometry. As a result, non-linear power behavior occurs in rapidly changing atmospheric conditions.

The conversion of irradiance to electrical output is also influenced by the temperature of the semiconductor. A rise in module temperature causes a drop in open-switch voltage, which results in an efficiency decline at midday that is noticeable despite the high radiation. These unfavorable conditions mainly affect fixed-tilt rooftop solar panels that are poorly cooled. Consequently, PV models have negative temperature coefficients included in them, and this making hourly performance dependent on the interaction of irradiance and temperature rather than on irradiance alone.

The principal factor that causes intermittency is the short-term variation of the atmosphere, and primarily the movement of clouds. The motion of the clouds can result in a decrease in the solar radiation within a matter of minutes, thus creating quick power-ups and downs, which in turn have an influence on grid voltage stability and inverter ramp-rate controls. Previous irradiance models have approached cloud cover as a simple binary input; however, newer methods have adopted cloud optical thickness and cloud-index variables. By directly taking into account the clouds' attenuation effects, Yang et al. achieved a significant improvement in hourly predictions through the use of cloud-index variables [8].

On a large scale, the photovoltaic system generation comes down to the sum of irradiance, PV conversion, DC–AC inversion, and grid interaction. The behavior of the inverter in addition to irradiance variations determines the amount of energy that can be used and hence delivered to the grid. State-of-the-art inverters not only allow for maximum-power-point tracking but may also apply thermal derating. The performance ratio metrics provide a standard measure for identifying the theoretical and actual energy delivered gap, thus allowing for a comparison of the performance across different climatic conditions.

From the forecasting point of view, the behavior of the sun consists of both deterministic components (solar geometry) and stochastic influences (clouds, aerosols, humidity). The classical methods would use autoregressive time-series methods to model irradiance, while the new ones would prefer to go for nonlinear statistical and hybrid learning algorithms. Nicoletti and Bevilacqua indicated that numerical weather prediction largely increases the accuracy of neural forecasting by involving atmospheric evolution [9].

Satellite observation increases the forecasting capability even more by allowing the estimation of irradiance in the locations without ground sensors. National-scale assessment gets its radiation data from satellite-derived hourly and long-duration datasets support the planning of future grid integration. Pfenninger and Staffell revealed that the analysis of multi-decadal hourly datasets gives a very clear understanding of long-term variability and regional resource behavior [10].

All in all, both theoretical and empirical evidence suggest that the hourly solar variation cannot be determined from daily or monthly averages because of different factors like the complex nature of the atmosphere, thermal effects, and inverter dynamics. Thus, modern research incorporates physical radiation modeling, numerical meteorology, neural learning, and real-time sensor analytics to precisely characterize photovoltaic generation at operational timescales.

## III.    RELATED WORK

The progress of hourly solar energy analysis research has come through various methodological changes first starting with empirical modeling of solar radiation and slowly but surely, reaching the point of data-driven and machine-learning approaches. One of the major contributions during this transition was the study of Jiménez-Pérez and Mora-López, who revealed that clustering and classification techniques could substantially enhance the forecast of hourly global solar radiation by dividing the weather conditions into separate irradiance regimes instead of treating them all as having a uniform daily behavior [1]. Their approach has shed light on the drawbacks of the simple averaging techniques and has demonstrated that the sub-daily variability can be captured only when the classification strategies take into account the cloud behavior and the irradiance transitions. Their results facilitated the high-resolution temporal analysis to be considered as an important factor for the future prediction of the forecasting applications and led to the improvement of the methods that would be able to cope with the fast-changing atmospheric conditions.

Contemporary studies have gradually followed this line of thought and have turned to the use of ensemble methods and short-term forecasting techniques namely the ones used for solar power prediction. Pan and Tan made use of day-ahead forecasting and hybrid model, which was a combination of clustering and ensemble approaches. Their results demonstrated that hybrid model architecture could beat individual neural or statistical models, particularly when the weather conditions quickly changed within a single-day [6]. The merit of such contribution is that it has proven ensemble-based learning to be less sensitive to model choice and thus, more reliable for grid-oriented forecasting, which makes the difference between hourly prediction research and periodic resource assessment based on monthly or yearly irradiance values.

The aforementioned developments were paralleled by numerous methods which incorporated atmospheric features directly into the forecasting architectures with emphasis on cloud-related indices. The researchers Yang et al. suggested a new technique for irradiance modeling that unified time-series decomposition and a cloud-cover index. This was done by showing that irradiance transitions are largely dependent on short-lived atmospheric conditions instead of long-term radiative patterns [8]. The method supplied proof that irradiance ramping and shading effects needed to be captured at high temporal frequency for the reliable PV power forecasting to be done. The researchers' outcomes led to the transition from purely statistical forecasting to hybrid physical-statistical models which are capable of including meteorological influence in near real time[12].

Deep learning has been a mainstay in modern forecasting methods, however, the early neural network applications had already identified the importance of nonlinear mapping for diffuse radiation estimation. Soares et al. presented neural networks as effective tools for the hourly diffuse radiation modeling in the urban scenario and indicated that the atmosphere in complex environments created variability that couldn't be treated with traditional regression-based models alone [16]. Their work is still pertinent since most of the current deep learning designs are still mechanized by the basic observation that nonlinear atmospheric effects impact irradiance conversion in a way that demands learning-based rather than analytical approximations.

Together, these different studies contributed to the clarification of a methodological shift from the empirical solar modeling to the data-driven forecasting. In past research, the focus was mainly on the estimation of irradiance and the recognition of radiation patterns but the latest research is drawing in atmospheric variability, ensemble forecasting, and cloud-sensitive learning to support operational decisions at hourly resolution. Even then, the different methods they have used, the studies all come to the same fundamental point: that the prediction of solar energy on an hourly basis does require computational models that can portray the nonlinear environmental processes and the short-term atmospheric dynamics which are not captured by daily or monthly averaging methods.

**Table 1 Summary of Selected Studies Related to Hourly Solar Analysis**

| Study | Focus Area | Data Type | Methodology | Main Contribution |
|---|---|---|---|---|
| Jiménez-Pérez & Mora-López (2016) [1] | Hourly radiation modeling | Ground irradiance | Clustering + classification | Improved hourly forecast accuracy under variable sky conditions |
| Vijayakumar et al. (2005) [2] | Short-term atmospheric variation | Ground radiation | Statistical analysis | Demonstrated significance of short-term radiation fluctuations |
| Ma et al. (2017) [3] | Intelligent energy networks | Sensor + grid data | Data analytics | Introduced data analytics role in smart grid operation |
| Yasmeen et al. (2022) [4] | Solar adoption impact | National datasets | Econometric analysis | Linked solar capacity, energy investment, and environmental benefit |
| Pan & Tan (2019) [6] | Day-ahead hourly PV forecasting | Measured irradiance | Cluster analysis + ensemble | Proved ensemble models outperform single ML algorithms |
| Yang et al. (2012) [8] | Cloud-sensitive forecasting | Cloud index | Statistical + cloud index | Introduced cloud-cover-based hour-level forecasting |

| Soares et al. (2004) [16] | Urban diffuse radiation | Local irradiance | Neural networks | Modeled diffuse radiation under complex urban conditions |
|---|---|---|---|---|

## IV.    METHODS AND FORECASTING MODELS

The development of photovoltaic operation and the increased demand for grid management in renewable-integrated power systems lead to the considerable evolution of hourly solar energy forecasting. Classic stochastic estimation was one of the earliest methods to solar radiation modeling, where mainly short-term statistical variations and persistence-based projections characterized the modeling approach. The methods assumed future irradiance to be very closely reflecting the recent historical values and thus performed well under clear sky conditions but were worsened greatly in places experiencing rapid cloud movements or high short-term atmospheric variability. Later on, researchers developed the situation by adding short-range empirical corrections based on meteorological measurements, mainly temperature and humidity, which were found to be the most important factors causing hourly radiation fluctuation in high-irradiance environments [2].

With the arrival of sensor-activated solar datasets, it became possible to implement the forecasting methodology in accordance with actual photovoltaic performance, and to gradually replace the analytical radiation methods with data-analytic ones. More importantly, intelligent energy network research pointed out that forecasting could no longer be regarded as solely dependent on environmental factors but required close collaboration between operational grid behavior, automated decision support, and energy-aware control mechanisms [3]. This change in thinking was at the heart of the transition from conventional irradiance modeling to real-time forecasting. Nowadays, models are more often than not driven by data, which consists of input from on-site pyranometers, satellite cloud estimation, sensor networks, and supervisory inverter monitoring. Given that irradiance can change quite drastically in a matter of minutes, the precision of forecasting is nowadays more and more tied to the use of temporal processing architectures that can seize very rapid atmospheric changes.

The application of machine learning techniques has resulted in significantly increased forecasting capacity through the discovery of non-linear connections between the inputs, i.e. the meteorological conditions, irradiance trends, and photovoltaic power generation. The use of support vector regression and gradient boosting methods has shown very noticeable improvements when compared with conventional statistical analysis, especially if large multi-season datasets are used for model training [4]. Machine-learning systems bring to light latent interactions that very often go unnoticed in physical irradiance equations, especially those produced by cloud movement, air mass shift, and regional climate behavior. These advantages facilitated rapid adoption of machine learning in industry-scale photovoltaic decision workflows and encouraged deployment of optimized inverter control, predictive maintenance, and energy-market participation based on short-term resource availability.

Although Artificial Intelligence is increasingly preferred for contemporary forecasts, hybrid models are significant in photovoltaic forecasting as physical processes governing irradiance forecast renders them unrepresentable only by statistical learning. Early techniques of clustering established that one should first derive the multiple atmospheric regimes before employing any forecasting models. Then, the idea was extended with the introduction of ensemble methods that bring many fragile learners together to generally minimize prediction using abrupt meteorological transitions [5]. Eventually, significantly enhancing the role of satellite-driven irradiance retrieval was able to extend forecasting scope strongly into wide solar fields and hereby entitled to demonstrate predictions in regions with minimal local sensor availability. Through this approach, the emphasis on spatial variation in accounting for the short-term photovoltaic output was reinforced while encouraging national-scale forecasting development.

Sub-hour data sets emerged as a methodological watershed for remote-sensing observations coupled with dynamic solar radiation analysis over the entire solar domain. This action enhanced the capacity to simulate high-resolution behavior of irradiance as it fluctuates through thousands of kilometers and also to demonstrate the photovoltaic variability at the utility scale [7]. The integration of sub-hour satellite data and machine learning frameworks significantly elevated predictive accuracy, especially within regions characterized by highly dynamic cloud behavior. In summary, sensor-based irradiance, remote-sensing technology, and neural processing reign together and provide hybrid frameworks, which bridge the actual atmospheric estimation and statistical learning efforts.

Deep learning methods introduced the long-sequence temporal representation by treating irradiance fluctuations as a stochastic time series rather than as independent snapshots. So, meteorological forecasting was regarded as a sequentially dynamic system, and artificial neural networks were modified in such a way that they could access memory effects reaching further back than just the short intervals. Numerical weather prediction—an input feature for neural networks—proved to be a great enhancer of the forecasting performance in the operational photovoltaic environment since it not only tied the atmospheric evolution but also allowed longer forecast horizons [9]. These architectures are capable of not just propagating the uncertainty, but also of learning the seasonal transformation and capturing the multicolored irradiance variations caused by the different layers of the cloud fields.

Hybrid forecasting models have been developed and are currently combining deep learning, satellite imaging, physical irradiance constraints, and numerical weather estimation. A typical architecture would include convolutional processing for cloud detection, recurrent memory for hourly radiation output, and ensemble decision layers for grid-oriented predictive control. Forecasting accuracy does, however, deteriorate under seasonal extremes and high-latitude conditions; hence, contemporary models have integrated uncertainty adjustment, multi-input transformation, and adaptive learning mechanisms, thus ensuring consistent forecast performance no matter the seasonal irradiance change [13].

On the operational level, forecasting models are becoming more and more influential in determining the dispatching of grid-connected photovoltaics, the scheduling of storage and inverter derating strategies. The forecast output is widely used in demand response activities, reserve allocation, and curtailment decisions, especially in those areas where solar has already reached or is very close to the level of penetration that traditional baseload plants can handle. Thus, photovoltaic forecasting is no longer seen as a meteorological modeling problem, but rather as a multidisciplinary research field that combines climate science, machine learning, power electronics, grid control, and economic optimization.

## V.    DATASETS FOR HOURLY SOLAR ANALYSIS

Hourly solar analytics are progressively depending on significant data repositories that will not only contain irradiance but also other factors like environmental parameters and PV performance of the system operated at such a high temporal resolution that it would be able to support forecasting and performance evaluation. In the past, the mainly used ground-based pyranometer measurements were the main source of irradiance information. However, recent activities have turned towards satellite-derived radiation grids and reconstructed measurement datasets that are capable of carrying out evaluations over multiple regions and over a longer period.

One of the most remarkable advancements that made it possible to do so was the work of Hummon et al. who created sub-hour solar data made from static spatial variability analysis and remote sensing information, which allowed the assessment of utility-scale irradiance and characterization of short-term variability over large areas [7]. These datasets; in addition to their providing temporal granularity suitable for operational simulation, have become the most common datasets in power-system modelling studies.

At the same time, Lamagna et al. proved that even when only monthly billing or limited sensor observations are available, hourly load and solar generation profiles can be reconstructed, thus showing that hourly curves can be mathematically inferred when direct measurements are missing [13]. This method is of great importance for developing countries where the use of continuous measurement devices and long-duration monitoring infrastructures is still very limited.

In order to make global comparisons and facilitate international comparison, open-access global platforms have also been introduced. McPherson et al. presented a web-based instrument powered by NASA MERRA reanalysis data that allows for the direct access of the hourly wind and solar PV generation numbers for different continents thus allowing for research and power-system assessment without any boundaries [17]. Such platforms provide the means of conducting hourly energy analysis at very low costs by offering non-isolated, time-consistent radiation series that are distributed freely.

Not long ago, Campbell et al. also published a coincidental wind-solar dataset for the continental United States which covered a long period and had long-term hourly values as well as a quality intended to support resource adequacy, reliability assessment, and renewable planning studies [21]. The dataset allows for the evaluation of both solar and wind scenario and is a pioneer in hybrid renewable resource scheduling research.

These datasets, which are in one way or another related, say that the data coming from isolated ground-sensor records are being replaced by global, satellite-supported, multi-decadal repositories that allow for forecasting, operational analysis, and large-scale renewable energy modeling. As the demand for hourly solar analytics keeps on growing, access to standardized multi-parameter datasets will become even more important for regional comparison, forecasting robustness, and real-time energy-system optimization.

## VI.    ENVIRONMENTAL INFLUENCES AND INVERTER PERFORMANCE BEHAVIOR

The hourly performance of photovoltaic systems is basically dependent on the interplay among the environmental factors, the dynamic temperature of the semiconductors, and the behavior of the inverters during the conversion process. The sunlight or irradiance is the most significant factor contributing to the power generated by a photovoltaic cell, however, more and more studies reveal that the effects created by thermal, humidity and regional atmospheric loading stresses are secondary but when added up over a long period of time, they can greatly influence the operational efficiency. The main challenge in temporal resolution of photovoltaic output lies in the highly dynamic nature of irradiance, which varies quickly over very short time intervals and which might cause sudden changes that are often overlooked by conventional averaging. It can be inferred from this complexity that temperature is a major factor determining the conversion efficiency because the voltage of the semiconductor junction decreases when the solar panel gets hotter, consequently changing the maximum power characteristics at the inverter stage. Past empirical studies assessed seasonal irradiance fluctuations in terms of static resource evaluation, but up-to-date data sets indicate that temperature variation is a major contributor to hourly efficiency reduction in particular during high-irradiance periods in tropical and arid regions [11].

Hourly ambient temperature rises throughout the day, usually following the peaks of the irradiance by a few hours, while the module temperature rises pretty much in line with the incoming radiation as a result of the absorption of the short-wave energy by the photovoltaic layers. In the case when the heat in the solar panel gets too high, the module temperature may reach as much as twenty degrees Celsius higher than the surrounding air. The voltage drop at the maximum power point during these times, when the irradiance is the highest and the temperature is already at its peak, indicates the predominant electrical effect of such thermal excess in the afternoon hours. Therefore, the power output does not necessarily follow the irradiance strictly but rather reflects the nonlinear interactions of thermal coefficients, material characteristics, and inverter limitations. The findings of regional studies in the high-irradiance areas showed that those places with the richest solar resource might have a larger thermal loss at the same time; consequently, the photovoltaic systems in these areas need to adopt better thermal control strategies that include the use of ventilation, convective cooling, and reflective coating technologies [14].

The synchronicity of irradiance and inverter temperature with respect to time results in an operational behavior that is usually referred to as afternoon efficiency roll-off. The panel temperature during the morning hours is moderate so that the recovery of the efficiency of the photovoltaic panels is rapid as irradiance is also increasing. Nevertheless, after the module temperatures have reached the upper limit of their nominal operating range, the efficiency starts to be reduced very slowly even though the irradiance is still at a high level, hence there will be a disparity between the solar energy that is available and the actual electrical output. This effect is most clearly seen in the yearly datasets that are obtained from the areas that have a lot of summer radiation and very little wind for cooling. The thermal effects are further exacerbated by the accumulation of dust, changes in humidity, and the dynamics of cloud reflectance, thus a relationship of coupling between the meteorological factors and photovoltaic conversion is established.

Hourly datasets have also shown that there are vast differences in thermal sensitivity between different inverter models and photovoltaic technologies. The semiconductor layers employed in monocrystalline silicons typically have lower temperature coefficients than those used in thin-film technologies, thus incurring less high-temperature derating. Inverters, likewise, have maximum temperature limits at which DC-to-AC conversion becomes less efficient or derating modes are automatically engaged to lower the thermal load. Hence, monitoring at the inverter level is crucial for determining the operational limits and for diagnosing the possible overheating effects that might lead to short-term output loss. Continuous operational measurements from campus-scale energy networks have shown that the hourly environmental variation not only affects the grid-side energy consumption but also reflects the interdependence of generation and local demand cycles [18].

In dusty and coastal climates, hourly radiation attenuation is strongly influenced by humidity, atmospheric aerosols, and micro-climate variability. The presence of aerosols in the atmosphere causes the direct beam irradiance to decrease, and at the same time, the diffuse components to increase, which makes it difficult to predict the irradiance patterns through the use of classical clear-sky formulas. Therefore, high-frequency atmospheric fluctuations may introduce more uncertainty than daily climate averages indicate. Similar results have been observed in temperate regions where it has been shown that cloud opacity is not a strict function of the atmospheric temperature. Consequently, cloud-induced irradiance loss can take place even during moderate thermal phases if high-altitude moisture layers are present. Such occurrences heighten the need for data-driven forecasting methods that use several atmospheric variables as input rather than relying solely on irradiance estimates.

Inverter performance are determined not only by their thermal fluctuations but also by the gradual effects of the environment, semiconductor aging, and grid-integration limits. The performing efficiency is reliant on both the instant DC voltage and the internal inverter electronics which, in turn, compensate for thermal stress by changing the switching pattern. When it comes to the areas with high PV penetration, the grid dispatching rules may demand curtailment during the peak irradiance times to maintain the system balance, hence, the maximum output of energy cannot always be achieved even if there are no thermal losses. Hence, it becomes essential to consider the efficiency data of inverters in the larger context of grid demand, scheduling, and dispatching procedures. Particularly, the areas with thick solar installation coverage are more and more subject to grid-induced curtailment which conceals the potential performance of the photovoltaics, and thus the apparent conversion efficiency is lowered despite the fact that the module-level output is still theoretically high.

Wind speed, though it has a natural cooling effect, can change a lot from one micro-climatic zone to another, which means that the hourly thermal dissipation would be different a lot depending on the area. According to weather reports, small-scale differences in the land might lead to the creation of wind shadows and the area around the rooftop heating up, especially in places where the rooftop is surrounded by urban heat sinks. These findings have made it clear that inverter analysis on an hourly basis should be done while taking into account the micro-geography and not the bulk environmental averages only. For example, a large-scale climatological study of wind and solar generation across Germany showed this interaction between cloud systems, circulation patterns in the region, and the power output at hourly resolution. It also pointed out how the use of regional atmospheric classification frameworks can highlight short-term photovoltaic variability that would otherwise be lost in daily aggregates [19].

The environmental behavior at hourly resolution is a strong argument in favor of using high-frequency monitoring, data fusion of both satellite and ground measurements, and sensor-based characterization of module temperature dynamics. Such environmental analysis would not only provide operators with necessary information related to operational stability, inverter longevity, and energy-yield forecasting but also help in the identification of critical time intervals when thermal derating could result in energy losses even under the irradiation conditions considered favorable.

## VII.    RESEARCH GAPS AND CHALLENGES

In spite of the considerable advances talked about in the previous paragraphs, there are still some limitations in research areas like solar energy hour analysis which result from the interaction of environmental variables, the accuracy of predictions, long-period inverter monitoring, and the lack of multi-season datasets. Most of today's research is mainly concerned with forecasting the amount of solar radiation or the power output of photovoltaic cells, while only a very small part of it tries to realistically measure the whole causal chain of the interaction between radiation, heat, and the electric making through the normal operation of the talk. Numerous studies spotlight accurate short-range forecasts from the application of machine-learning but these models are often trained on datasets that are limited to a few climates which makes it hard to apply them to other climates where the seasonal or regional dynamics are different from that of the training distributions [15].

Not only limited to short periods, most papers tend to indicate forecasting performance as the main focus leaving inverter deterioration and loss of efficiency over time unnoticed. Powering up every hour shows that heating up and cooling down, switching stress, and exposure to the elements all together eventually weaken the inverter parts. Still, the long-term analysis of inverters' performance has been limited as the researchers usually do not have access to uninterrupted and high-quality operational datasets. Because of the systematic degradation measurements that are usually lacking, it then becomes impossible to accurately model the life cycle and the less developing predictive maintenance. The works of Campbell et al. demonstrate how important it is to have multi-decadal operational data

for analyzing the behavior of renewables but a similar case of high-resolution time span for inverter-level photovoltaic performance is very rare in open access literature [21].

A clear case of methodological fragmentation is presented by the forecasting literature. Conducting comparative studies of the methods is still very limited in the case of support vector regressors, ensemble learning, and recent neural networks, and all of these methods showing great predictive power at the same time. The majority of the articles assess forecasting models on the basis of datasets taken in isolation and seldom carry out strict benchmarking across different locations. This hinders the researchers to decide whether such improvements in forecasting are due to the model design or simply reflect the conditions in the areas concerned. What's more, the models published usually consider the spatial-temporal characteristics to be the same, where in fact the operational irradiance is changing quite a lot among the different climate areas, which indicates the need of a forecasting approach that is especially adaptive to differences in seasons, regions and atmospheres.

Another important challenge for research comes from the unavailability of environmental measurements. A number of datasets contain data on the irradiance only but do not count on humidity, aerosol thickness, wind speed and dust accumulation. Therefore, the forecasting models often have to make assumptions about the power-output behavior based on irradiance, and this in turn may lead to significant underestimate of the influence of diffuse radiation and thermal properties. The thermal coefficients are sometimes made available at the module level but hardly ever at the inverter level, even though the inverter manufacturers are equipped with the technology that allows them to non-invasively monitor temperature fluctuations within the unit. Thus, the thermal sensitivity has been extensively studied from the viewpoint of photovoltaic materials but is still quite less defined in the context of the inverter electronics and long-term operational reliability.

One more ongoing disparity is the lack of in-depth investigation into microgrid and residential solar systems distributed installations, which, however, is a trend being witnessed mostly in renewable energy literature, marking utility-scale systems and ground-mounted solar farms as priority areas. This is partly due to the fact that utility installations have advanced monitoring systems compared to rooftop systems which present different thermal behavior due to limited ventilation, urban heat effects, and increased dust deposition. Energy analytics for residential deployments are overlooked even though they have larger hourly fluctuation, more diverse systems, and more dynamic load interaction. Predictive analytics and adaptive inverter control can support these systems a lot, especially in less developed regions where grid conditions are still risky and storage infrastructures are just being developed.

More and more difficulties are to come from the increase in renewable penetration and the new interaction patterns between solar energy and electric grids. In cases of high generation, grid operators exercise curtailment or impose reactive power limits to maintain the grid's stability, hence the solar performance being more dependent on grid's operational behavior rather than on irradiance. Consequently, operational curtailment leads to less energy yield being witnessed even though the environmental potential is high. This creates a conceptual problem because forecasting tools based on historical data might consider curtailment-induced output reduction as system failures instead of recognizing it as a consequence of grid dispatch protocols. Therefore, in such cases, hourly analysis has to separate environmental behavior from grid-side actions to get a true picture of photovoltaic potential.

forecasting tools based on historical data might consider curtailment-induced output reduction as system failures instead of recognizing it as a consequence of grid dispatch protocols. Therefore, in such cases, hourly analysis has to separate environmental behavior from grid-side actions to get a true picture of photovoltaic potential.

The limitations mentioned above suggest that solar hourly analysis has made a notable progress in the forecasting area but still needs development in terms of environmental integration, long-duration operational assessment, and multi-regional generalization. The literature is very clear that there is a high analytical potential but at the same time it needs larger datasets, more in-depth inverter lifecycle studies, and internationally synchronized data infrastructures in order to completely express hourly photovoltaic behavior over different climates, seasons, and deployment scenarios.

## VIII.     CONCLUSION

This review took a look at solar energy characteristics on an hourly basis based on recent analytical and performance-oriented studies. The literature indicates that hourly variation of irradiance is still the main factor affecting the output power of PV systems, while temperature has a secondary but steady contribution to the conversion efficiency and inverter performance. The transition from daily means to high-resolution datasets has helped greatly in understanding short-lived phenomena like cloud-induced intermittency, thermal effects, and inverter efficiency degradation. Moreover, hourly measurement has become indispensable for operational tasks such as grid integration, storage scheduling, and predictive maintenance planning.

Incorporating machine learning, satellite sensing, and neural forecasting is the recent research that has stepped beyond conventional statistical methods to improve short-term prediction. These models allow for the inverter loading to be planned more accurately and further showcase the merit of hybrid physical–data-driven approaches for the operational decision-making process. However, there are still significant challenges like the unavailability of unified environmental datasets, lack of long-term monitoring of inverters, and no multi-site comparative evaluation.

Besides environmental conditions, the photovoltaic performance is increasingly influenced by grid-side behavior like curtailment, regulation, and demand balancing, as the review points out. With the increase in solar penetration, the forecasting tools will have to take into account not only atmospheric variability but also operational grid constraints. On the other hand, distributed and residential systems form a gap in the knowledge pool because of their thermal behavior and local demand interaction being entirely different from the grid.

The future advancement of renewable energy technologies will rely heavily on the eventualities of the long-term datasets that will be formed from the parameters of irradiation, temperature, and inverter performance at different places and varied weather conditions along with the forecasting models that will be able to accommodate seasonal changes and climate-driven shifts in environmental conditions. The maximum hourly solar research period is now encroaching into meteorology, power electronics, and artificial intelligence areas, thus creating opportunities for more accurate forecasting and more intelligent system control. The continuous development of machine learning, sensor technologies, and monitoring infrastructure will not only enhance the hourly solar analytics but will also be a great support to the global renewable energy transition endeavors.

## REFERENCES

[1]     P. F. Jiménez-Pérez and L. Mora-López, "Modeling and forecasting hourly global solar radiation using clustering and classification techniques," Solar Energy, vol. 135, pp. 682–691, 2016.

[2]     G. Vijayakumar et al., "Analysis of short-term solar radiation data," Solar Energy, vol. 79, no. 5, pp. 495–504, 2005.

[3]     Z. Ma, et al., "The role of data analysis in the development of intelligent energy networks," IEEE Network, vol. 31, no. 5, pp. 88–95, 2017.

[4]     R. Yasmeen, et al., "Exploring the role of solar energy and foreign direct investment for clean environment: Evidence from top 10 solar energy consuming countries," Renewable Energy, vol. 185, pp. 147–158, 2022.

[5]     Jiménez-Pérez, Pedro F., and Llanos Mora-López. "Modeling and forecasting hourly global solar radiation using clustering and classification techniques." Solar Energy 135 (2016): 682-691.

[6]      Pan, Cheng, and Jie Tan. "Day-ahead hourly forecasting of solar generation based on cluster analysis and ensemble model." IEEE Access 7 (2019): 112921-112930.

[7]      Hummon, Marissa, et al. Sub-hour solar data for power system modeling from static spatial variability analysis. No. NREL/CP-6A20-56204. National Renewable Energy Lab.(NREL), Golden, CO (United States), 2012.

[8]      Yang, Dazhi, Panida Jirutitijaroen, and Wilfred M. Walsh. "Hourly solar irradiance time series forecasting using cloud cover index." Solar Energy 86.12 (2012): 3531-3543.

[9]      Nicoletti, Francesco, and Piero Bevilacqua. "Hourly photovoltaic production prediction using numerical weather data and neural networks for solar energy decision support." Energies 17.2 (2024): 466.

[10]      Pfenninger, Stefan, and Iain Staffell. "Long-term patterns of European PV output using 30 years of validated hourly reanalysis and satellite data." Energy 114 (2016): 1251-1265.

[11]      Armstrong, S., and W. G. Hurley. "A new methodology to optimise solar energy extraction under cloudy conditions." Renewable energy 35.4 (2010): 780-787.

[12]      Suh, Jangwon, and Yosoon Choi. "Methods for converting monthly total irradiance data into hourly data to estimate electric power production from photovoltaic systems: A comparative study." Sustainability 9.7 (2017): 1234.

[13]      Lamagna, Mario, et al. "Hourly energy profile determination technique from monthly energy bills." Building Simulation. Vol. 13. No. 6. Beijing: Tsinghua University Press, 2020.

[14]      Awan, Ahmed Bilal, et al. "Solar energy resource analysis and evaluation of photovoltaic system performance in various regions of Saudi Arabia." Sustainability 10.4 (2018): 1129.

[15]      Giannuzzo, Lorenzo, et al. "Reconstructing hourly residential electrical load profiles for renewable energy communities using non-intrusive machine learning techniques." Energy and AI 15 (2024): 100329.

[16]      Soares, Jacyra, et al. "Modeling hourly diffuse solar-radiation in the city of São Paulo using a neural-network technique." Applied energy 79.2 (2004): 201-214.

[17]      McPherson, Madeleine, et al. "An open-access web-based tool to access global, hourly wind and solar PV generation time-series derived from the MERRA reanalysis dataset." Energies 10.7 (2017): 1007.

[18]      Liao, Wei, et al. "A twenty-year dataset of hourly energy generation and consumption from district campus building energy systems." Scientific data 11.1 (2024): 1400.

[19]      Drücke, Jaqueline, et al. "Climatological analysis of solar and wind energy in Germany using the Grosswetterlagen classification." Renewable Energy 164 (2021): 1254-1266.

[20]      Xu, Yaojian, et al. "A complementary fused method using GRU and XGBoost models for long-term solar energy hourly forecasting." Expert Systems with Applications 254 (2024): 124286.

[21]      Campbell, Allison M., et al. "A Multi-Decadal Hourly Coincident Wind and Solar Power Production Dataset for the Contiguous United States." Scientific Data 11.1 (2024): 1121.